(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) EP 0 986 046 A1

# (12) EUROPEAN PATENT APPLICATION

(43) Date of publication:
**15.03.2000 Bulletin 2000/11**

(51) Int. Cl.[7]: **G10H 1/00**, G10H 3/12

(21) Application number: **99306932.7**

(22) Date of filing: **31.08.1999**

(84) Designated Contracting States:
**AT BE CH CY DE DK ES FI FR GB GR IE IT LI LU MC NL PT SE**
Designated Extension States:
**AL LT LV MK RO SI**

(30) Priority: **10.09.1998 US 150582**

(71) Applicant: **LUCENT TECHNOLOGIES INC.**
**Murray Hill, New Jersey 07974-0636 (US)**

(72) Inventor: **Curtin, Steven D.**
**Freehold, New Jersey 07728 (US)**

(74) Representative:
**Watts, Christopher Malcolm Kelway, Dr. et al**
**Lucent Technologies (UK) Ltd,**
**5 Mornington Road**
**Woodford Green Essex, IG8 0TU (GB)**

## (54) System and method for recording and synthesizing sound and infrastructure for distributing recordings for remote playback

(57) Systems and methods for recording and synthesizing sound in a resolution-independent manner and infrastructures for distributing resolution-independent recordings for remote playback. In one

*FIG. 1*

## Description

### Technical Field Of The Invention

**[0001]** The present invention is directed, in general, to sound recording and reproduction and, more specifically, to a resolution-independent system and method for making a reducing of a sound and later employing the record to synthesize the sound.

### Background Of The Invention

**[0002]** Current recording of musical instrument performances are either based on sampling the analog signal of the instrument or recording the gestures that are input to a controller. This leads to playback situations where the performance may only be edited in a fixed time domain for the case of sampling, or a performance recording that may only realistically record keyboard and percussion in the case of Musical Instrument Digital Interface (MIDI). Additive and spectral synthesis technologies break down musical performances into discrete notes as opposed to a continuous performance. Generally, the ability to change the tempo or the key and to synchronize a performance with external events during playback are often difficult to accomplish at reasonable cost and without unacceptable distortions.

**[0003]** The sound waveforms produced may be characterized by many parameters, including frequency and amplitude. Using Fourier analysis, sound waveforms may be represented in a frequency domain as a spectral frame, consisting of spectral components. The spectral frame contains the waveform's lowest, or fundamental, frequency, along with its harmonics (spectral components which occur at multiples of the fundamental frequency). Spectral components from string instruments and from vowels in speech typically occur at close to whole number multiples of the fundamental frequency, while spectral components from percussion instruments often occur at non-integral multiples of the fundamental frequency.

**[0004]** Current sound recordings have been seen to be typically sample rate dependent or suffer from other recording and playback characteristics that make modifications to the record difficult to accomplish at acceptable costs or distortion levels. Additionally, these limitations make the offering of current sound recordings very limited in format for playback selection. Radio stations offer a selection of recordings that may be programmed for many days into the future with only an occasional specific request capability allowed, usually by telephone. Even in the case of selected requests, the recording is completely fixed in format with respect to. tempo and key as well as its basic arrangement.

particular sampling frequency.

**[0011]** In one embodiment of the present invention, the system further includes a mapping circuit that applies a temporal quantization map to the record. Once the record is created, the present invention accommodates a wide range of conventional and later-developed sound manipulation techniques.

**[0012]** In one embodiment of the present invention, the system further includes an editor that modifies a selected one of a content and an order of the record. The editor allows still further manipulation of the sound once recorded.

**[0013]** In one embodiment of the present invention, the frame analyzer identifies the selected one of common spectra structures and common formant structures by Fourier analyzing the frames. Those skilled in the art are familiar with, in particular, fast Fourier transform techniques by which frequencies may be analyzed. The present invention is compatible with other conventional or later-developed spectrum analysis techniques, such as wavelets.

**[0014]** The present invention further provides infrastructures for distributing recordings for remote playback. One infrastructure includes: (1) a radio station having a recording database associated therewith, (2) a plurality of recordings contained within the recording database, each of the plurality of recordings including a selected one of common spectra structures and common formant structures corresponding thereto, (3) a request receiver, coupled to the recording database, that receives remote requests for ones of the plurality of recordings and (4) a transmitter, coupled to the recording database, that transmits the ones of the plurality of recordings in response to the requests.

**[0015]** The present invention therefore provides what amounts to "audio-on-demand" wherein formatted audio files are provided to remote "radios" to allow the remote "radios" to synthesize the audio in situ. Therefore, in one embodiment of the present invention, the infrastructure further includes a plurality of remote radios capable of receiving and digitally manipulating the ones of the plurality of recordings. The remote radios may comprise software that can be downloaded and executed on data processing and storage hardware to allow the ones of the plurality of recordings to be played. This infrastructure sharply contrasts with conventional analog AM or FM radio infrastructures in which remote radios simply demodulate and amplify received radio waves.

**[0016]** In one embodiment of the present invention, the transmitter broadcasts the ones of the plurality of recordings to receivers. Alternatively, the ones of the plurality of recordings may be addressed to individual remote "radios."

**[0017]** In one embodiment of the present invention, the ones of the pluralities of recordings are embodied in a plurality of bitstream files. The bitstream files contain data pertaining to the fundamental frequencies and the selected one as described above.

**[0018]** In one embodiment of the present invention, the recording database contains a record of the requests. This allows song popularity or advertisement dissemination to be tracked and accurate royalty payments to be

Detailed Description

**[0021]** Referring initially to FIGURE 1, illustrated is a block diagram 100 of a resolution-independent system for recording a musical instrument constructed according to the principles of the present invention. The resolution-independent system of block diagram 100 includes a sound source 110, a sampler 120, a frame generator 130, a frame analyzer 140, a first storage unit 150, a mapping circuit 160, an editor 170 and a second storage unit 180.

**[0022]** The present invention provides systems and methods for recording and synthesizing sound in a resolution-independent manner and infrastructures for distributing resolution-independent recordings for remote playback. In the present embodiment, a musical instrument may be used to generate the sound. Those skilled in the art are familiar with the formant content of certain musical instruments, such as string and wind instruments. Human voices also contain formants that may be captured and employed in later synthesis to be used in play-back of the recording. The present invention can operate with any sound, however.

**[0023]** The sampling, which may occur at 1 ms intervals, is accomplished by the sampler 120 before the fundamental frequencies are extracted by the frame generator 130. However, those skilled in the art will understand that the present invention is not limited to a particular sampling frequency or the use of a separate sampler as shown. The sampler 120 may be included as part of the frame generator 130. This embodiment allows the data comprising a sound source to be independent of the sampling rate or tempo desired. The frame generator 130 creates frames and extracts fundamental frequencies and spectral envelopes from the sound source 110 through the sampler 120. Then, the frame analyzer 130 identifies a selected one of common spectral structures and common formant structures in the frames and creates a record, containing this selected one and the appropriate fundamental frequencies, that is then stored in the first storage unit 150.

**[0024]** The present invention therefore introduces the broad concept of storing fundamental frequencies and selected structures in sound and creating a record containing those fundamental frequencies and selected structures to provide a basis for subsequent synthesis to be used in playback. The present invention analyzes the sound as a continuous performance, irrespective of individual tones or notes. The frames may be discrete, allowing them to correspond to a discrete period of time. The frame analyzer 140 identifies the selected one of common spectral structures and common formant structures by Fourier analyzing the frames. Those skilled in the art are familiar with Fast Fourier Transform (FFT) techniques as one technique by which frequencies may be analyzed. The present invention is compatible with other conventional or later-developed spectrum analysis techniques, such as wavelets, as well.

**[0025]** In this embodiment, common spectral or formant structures may be contained in a dictionary in order to compress the total size of the record. The identification and grouping of common spectral or formant structures

in a step 215 before extracting the fundamental frequencies from the sampled sound signal. Frames, that may be discrete, are then generated in a step 220 from which fundamental frequencies and spectral envelopes are then extracted. These frames are then analyzed in a step 225 using Fourier analysis, and then common spectral and formant structures are identified in a step 230. These common structures are then stored in a step 235 which creates a record of the common structures. A temporal quantization map may then be applied to the record as shown in a step 240, or the record may be modified selecting one of a content and an order of the record in a step 245 in order to edit the contents as required. The method ends in an end step 250 where the sound has been selected, sampled, framed, analyzed, recorded, mapped or modified in this embodiment.

**[0030]** Turning now to FIGURE 3, illustrated is a block diagram 300 of a resolution-independent system for synthesizing a recorded musical instrument constructed according to the principles of the present invention. The resolution-independent system of block diagram 300 includes a storage unit 305, a mapping circuit 310, an editor 315, a waveshaper 320, an output device 325 and a speaker 330. The storage unit 305 contains the records and dictionaries that have been created by sampling, framing and analyzing the sound source as discussed in FIGURE 1 and FIGURE 2 above.

**[0031]** The mapping circuit 310 applies a temporal quantization map to the record. As stated earlier, the present invention may accommodate a wide range of conventional and later-developed sound manipulation techniques allowing the record to be synthesized with a changed tempo or key and provide the ability to synchronize the playback performance with an external clock. Further manipulation of the recorded sound is provided with the editor 315 which may modify the content or the order of the record to provide effects not contained in the original sound source. The editor 315 may also be used to re-arrange the record sequence in time relative to the frequency contour as discussed.

**[0032]** The waveshaper 320 is coupled to the storage unit 305, the mapping circuit 310 and the editor 315. The waveshaper 320 takes the fundamental frequencies and applies a waveshaping transfer function to create a waveform from either the stored record, the mapped record or the edited record. The waveshaper 320 may also use some combination of these three in order to generate the waveform. The waveshaper 320 may also select from a number of waveshaping transfer functions that are stored in the waveshaper 320 to accommodate the waveshaping process. The waveshaper 320 is clocked externally, in this embodiment, allowing the waveform to be synchronized with external events. The waveform may then be converted into an output sound using the output device 325 and the speaker 330. The synthesis process represented here allows the originally recorded sound to be reproduced with appropriate fidelity or allows the originally recorded sound to be modified as deemed appropriate to the user. This does not preclude the use of other synthesis techniques such as FFT or direct sine.

**[0033]** Turning now to FIGURE 4, illustrated is a flow diagram 400 of a resolution-independent method of synthesizing a recorded musical instrument that may be carried out in the system of FIGURE 3. The method depicted in

synthesize the audio in situ. In this embodiment of the present invention, the infrastructure further includes a plurality of remote digital radios capable of receiving and digitally manipulating the plurality of recordings. This infrastructure sharply contrasts with conventional analog AM or FM radio infrastructures in which remote radios simply demodulate and amplify received radio waves. This infrastructure may function with any currently-proposed or later-developed digital transmitter and receiver standards. The prograr material for the recordings may include but is not limited to weather, news, stock quotes or other topical information.

**[0039]** The random access playlist 515 represents the pluralities of recordings, which are embodied in a plurality of bitstream files. The bitstream files contain data pertaining to the fundamental frequencies and the selected one as described above. The bitstream files, which may represent a collection of selections or the collection of offerings, may occur in a single serial loop. The user may select the ones of these that are downloaded and played. Alternately, they may occur in a collection of parallel loops allowing the user to perform the download 520 more rapidly. The transmitter, associated with the wireless server 505, may broadcast the ones of the plurality of recordings to all remote radios. Alternatively, the ones of the plurality of recordings may be addressed only to individual remote radios. The recording database may contain a record of the requests. This allows song popularity or advertisement dissemination to be tracked and accurate royalty payments to be calculated automatically.

## Claims

1. A method of recording sound, for example sound produced by a musical instrument, comprising:

   extracting fundamental frequencies and spectral envelopes from said sound;

   creating frames from said fundamental frequencies and spectral envelopes;

   identifying a selected one of common spectra structures and common formant structures in said frames; and

   creating a record containing said fundamental frequencies and said selected one.

2. The method as recited in claim 1 wherein said frames are discrete.

3. A method as recited in claim 1 further comprising sampling said sound before extracting said

a transmitter, coupled to said recording database, that transmits said ones of said plurality of recordings in response to said requests.

9. The infrastructure as recited in claim 8 wherein said transmitter broadcasts said ones of said plurality of recordings to receivers.

10. The infrastructure as recited in claim 15 further comprising a plurality of remote radios capable of receiving and digitally manipulating said ones of said plurality of recordings.

11. The infrastructure as recited in any of claims 8 to 10 wherein said ones of said pluralities of recordings are embodied in a plurality of bitstream files.

12. The infrastructure as recited in any of claims 8 to 11 wherein said recording database contains a record of said requests.

13. A radio, comprising:

a receiver for receiving a recording including fundamental frequencies and a selected one of common spectra structures and common formant structures corresponding thereto;

a waveshaper, coupled to said receiver, for applying a waveshaping transfer function based on said selected one to said fundamental frequencies to create a waveform; and

a speaker, coupled to said waveshaper, for converting said waveform into an output sound.

14. The radio as recited in claim 13 further comprising a mapping circuit, coupled to said receiver, that applies a temporal quantization map to said selected one.

15. The radio as recited in claim 13 or claim 14 wherein said waveshaping transfer function is selected from a plurality of waveshaping transfer functions stored in said waveshaper.

FIG. 1

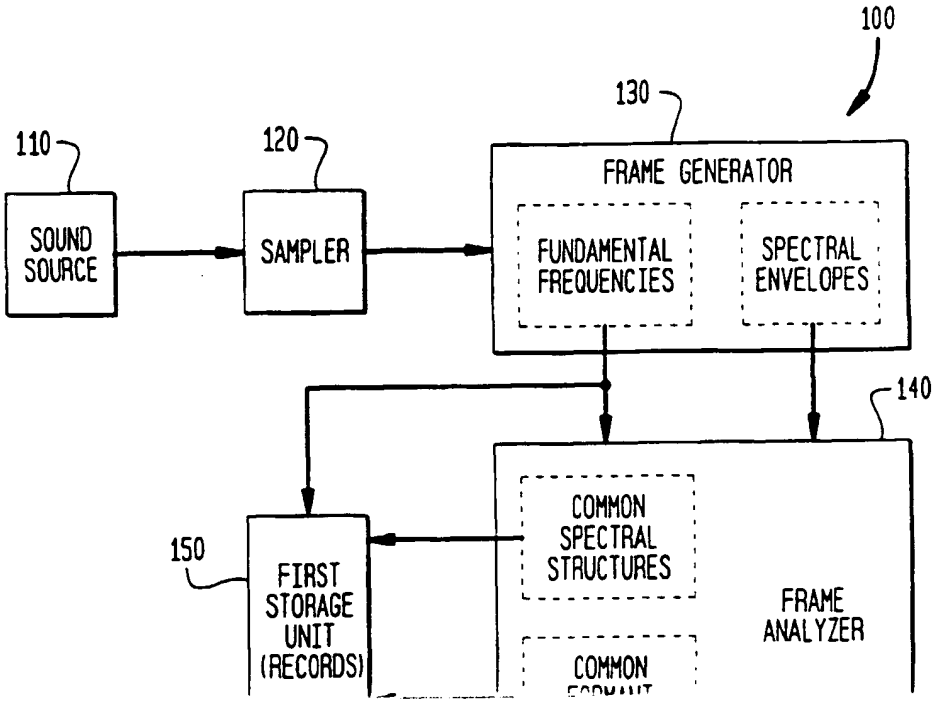

100
130
110
120
FRAME GENERATOR
SOUND SOURCE
SAMPLER
FUNDAMENTAL FREQUENCIES
SPECTRAL ENVELOPES
140
COMMON SPECTRAL STRUCTURES
150
FIRST STORAGE UNIT (RECORDS)
FRAME ANALYZER
COMMON

FIG. 2

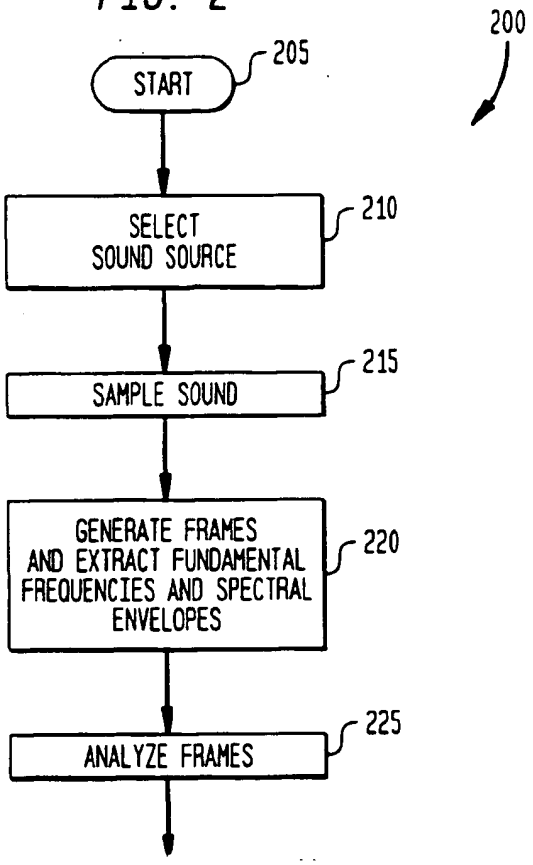

200
START
205
SELECT SOUND SOURCE
210
SAMPLE SOUND
215
GENERATE FRAMES AND EXTRACT FUNDAMENTAL FREQUENCIES AND SPECTRAL ENVELOPES
220
ANALYZE FRAMES
225

FIG. 3

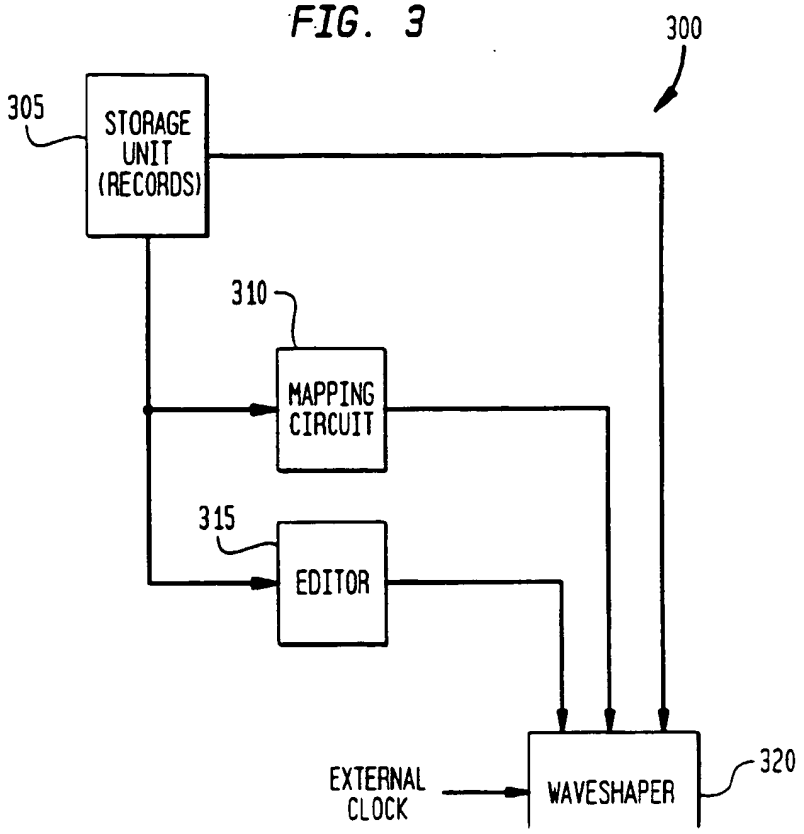

300
305
STORAGE UNIT (RECORDS)
310
MAPPING CIRCUIT
315
EDITOR
EXTERNAL CLOCK
WAVESHAPER
320

FIG. 4

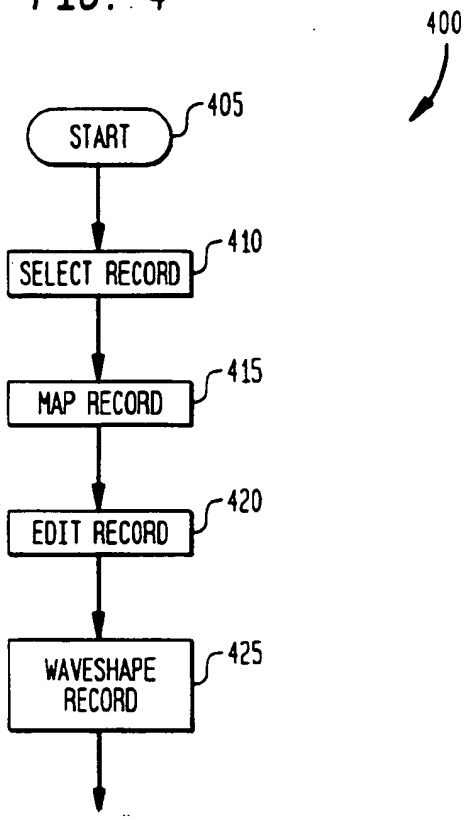
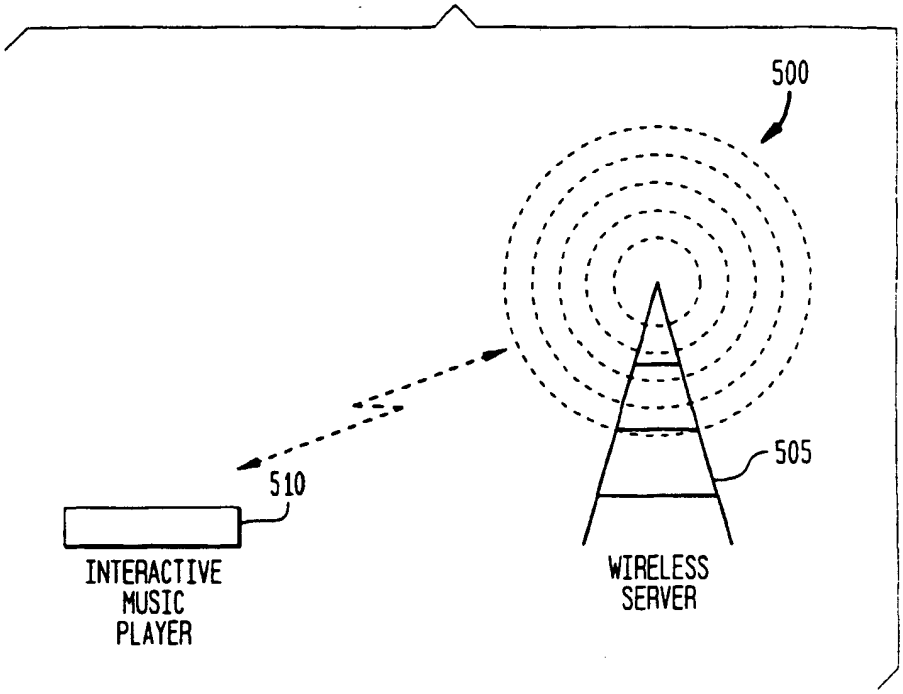

400
START
405
SELECT RECORD
410
MAP RECORD
415
EDIT RECORD
420
WAVESHAPE RECORD
425

FIG. 5A



European Patent Office

# EUROPEAN SEARCH REPORT

Application Number
EP 99 30 6932

## DOCUMENTS CONSIDERED TO BE RELEVANT

| Category | Citation of document with indication, where appropriate, of relevant passages | Relevant to claim | CLASSIFICATION OF THE APPLICATION (Int.Cl.7) |
|---|---|---|---|
| X | US 5 536 902 A (WOLD ERLING ET AL)<br>16 July 1996 (1996-07-16)<br>* column 3, line 7 - line 46 *<br>* column 9, line 58 - column 10, line 67; figure 2 * | 1-6 | G10H1/00<br>G10H3/12 |
| P,X | US 5 886 276 A (LEVINE SCOTT N ET AL)<br>23 March 1999 (1999-03-23)<br>* column 4, line 29 - column 6, line 67 *<br>* column 10, line 52 - column 11, line 5 *<br>* column 13, line 22 - column 14, line 25; figures 1,5 * | 1-17 | |
| A | EP 0 686 923 A (MOTOROLA INC)<br>13 December 1995 (1995-12-13)<br>* page 3, line 8 - line 34 * | 1,6,8 | |
| A | US 5 248 845 A (MASSIE DANA C ET AL)<br>28 September 1993 (1993-09-28)<br>* column 4, line 3 - line 21 *<br>* column 5, line 11 - column 6, line 10 *<br>* column 16, line 46 - line 68; figure 2 * | 1-7 | TECHNICAL FIELDS SEARCHED (Int.Cl.7) |
| A | US 5 735 744 A (OKAMOTO TAKEYA) | 8-17 | G10H |

## ANNEX TO THE EUROPEAN SEARCH REPORT ON EUROPEAN PATENT APPLICATION NO. EP 99 30 6932

This annex lists the patent family members relating to the patent documents cited in the above-mentioned European search report. The members are as contained in the European Patent Office EDP file on
The European Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

29-11-1999

| Patent document cited in search report | Publication date | Patent family member(s) | Publication date |
|---|---|---|---|
| US 5536902 A | 16-07-1996 | JP 2906970 B | 21-06-1999 |
| | | JP 7325583 A | 12-12-1995 |
| US 5886276 A | 23-03-1999 | NONE | |
| EP 0686923 A | 13-12-1995 | US 5473557 A | 05-12-1995 |
| | | JP 8006924 A | 12-01-1996 |
| US 5248845 A | 28-09-1993 | AU 3918293 A | 21-10-1993 |
| | | WO 9319455 A | 30-09-1993 |
| | | US 5698807 A | 16-12-1997 |
| US 5735744 A | 07-04-1998 | JP 2532198 B | 11-09-1996 |
| | | JP 6334780 A | 02-12-1994 |
| | | US 5775995 A | 07-07-1998 |
| | | US 5489103 A | 06-02-1996 |